Defence Science Journal, Vol. 58, No. 3, May 2008, pp. 338-352

Pixel-level Image Fusion using Wavelets and Principal Component Analysis

V.P.S. Naidu and J.R. Raol

National Aerospace Laboratories, Bangalore-160 017

ABSTRACT

Image registration and fusion are of great importance in defence and civilian sectors, e.g. ,recognising a ground/air force vehicle and medical imaging. Pixel-level image fusion using wavelets and principal component analysis has been implemented and demonstrated in PC MATLAB. Different performance metrics with and without reference image are implemented to evaluate the performance of image fusion algorithms. As expected, the simple averaging fusion algorithm shows degraded performance. The ringing tone presented in the fused image can be avoided using wavelets with shift invariant property. It has been concluded that image fusion using wavelets with higher level of decomposition showed better performance in some metrics and in other metrics, principal components analysis showed better performance.

Keywords: Pixel-level image fusion, wavelets transform, principal component analysis, multi-sensor image fusion

NOMENCLATURE

$\psi_{a,b}$	Mother wavelet
λ	Eigen value
μ_{I_r}	Mean of the reference image
σ	Standard deviation
cov	Covariance function
$CE(I_j;I_f)$	Cross entropy of the j^{th} input image and fused image
E	Expectation operator
$f(x)$	One dimensional signal
$h_{I_r I_f}$	Joint histogram of reference and fused images
(i,j)	Pixel index
I, I_f, I_L, I_H, I_r	Representation of source, fused, LPF, HPF, and reference images
$I_{LL}, I_{LH}, I_{HL}, I_{HH}$,	Representation of average, horizontal, vertical, diagonal information of I
I_1, I_2	Representation of input images to be fused
m,n	Integers
$norm$	Norm operator
M, N	Sizes of the image
P_1, P_2	Principal components
T	Matrix transpose
$CORR$	Correlation
CF	Column frequency of the image

Received 29 September 2006, revised 01 August 2007

He	Entropy
MAE	Mean absolute error
RF	Row frequency of the image
PFE	Percentage fit error
PSNR	Peak signal-to-noise ratio
RMSE	Root mean square error
SSIM	Measure of structural similarity

1. INTRODUCTION

Multi-sensor image fusion (MIF) is a technique to combine the registered images to increase the spatial resolution of acquired low detail multi-sensor images and preserving their spectral information. Of late MIF has emerged as a new and promising research area. The benefiting fields from MIF are: Military, remote sensing, machine vision, robotic, and medical imaging, etc. Some generic requirements could be imposed on the fusion scheme: (a) the fusion process should preserve all relevant information contained in the source images, (b) the fusion process should not introduce any artifacts or inconsistencies which would amuse the human observer or following processing stages, and (c) irrelevant features and noise should be suppressed to a maximum extent.

The problem that MIF tries to solve is to merge the information content from several images (or acquired from different imaging sensors) taken from the same scene in order to accomplish a fused image that contains the finest information coming from the original images[1]. Hence, the fused image would provide enhanced superiority image than any of the original source images. Dependent on the merging stage, MIF could be performed at three different levels viz. pixel level, feature level and decision level[2]. In this paper, pixel-level-based MIF is presented to represent a fusion process generating a single combined image containing an additional truthful description than individual source image.

The simplest MIF is to take the average of the grey level source images pixel by pixel. This technique would produce several undesired effects and reduced feature contrast. To over come this problem, multi-scale transforms, such as wavelets, Laplacian pyramids, morphological pyramid, and gradient pyramid have been proposed. Multi-resolution wavelet transforms could provide good localization in both spatial and frequency domains. Discrete wavelet transform would provide directional information in decomposition levels and contain unique information at different resolutions[3,4].

Principal component analysis (PCA) is a mathematical tool which transforms a number of correlated variables into a number of uncorrelated variables.The PCA is used extensively in image compression and image classification. Image fusion algorithm that utilises the PCA is described in this paper. The fusion is achieved by weighted average of images to be fused. The weights for each source image are obtained from the eigen vector corresponding to the largest eigenvalue of the covariance matrices of each source. Performance metrics are used to evaluate the wavelets, PCA, and simple average-based image fusion algorithms.

One of the important prerequisites to be able to apply fusion techniques to source images is the image registration, i.e., the information in the source images is needed to be adequately aligned and registered prior to fusion of the images. In this paper, it is assumed that the source images are already registered.

2. FUSION ALGORITHMS

The details of wavelets and PCA algorithm and their use in image fusion along with simple average fusion algorithm are described in this section.

2.1 Wavelet Transform

Wavelet theory is an extension of Fourier theory in many aspects and it is introduced as an alternative to the short-time Fourier transform (STFT). In Fourier theory, the signal is decomposed into sines and cosines but in wavelets the signal is projected on a set of wavelet functions. Fourier transform would provide good resolution in frequency domain and wavelet would provide good resolution in both time and frequency domains. Although the wavelet theory was introduced as a mathematical tool in 1980s, it has been extensively used in image processing

that provides a multi-resolution decomposition of an image in a biorthogonal basis and results in a non-redundant image representation. The basis are called wavelets and these are functions generated by translation and dilation of mother wavelet. In Fourier analysis the signal is decomposed into sine waves of different frequencies. In wavelet analysis the signal is decomposed into scaled (dilated or expanded) and shifted (translated) versions of the chosen mother wavelet or function . A wavelet as its name implies is a small wave that grows and decays essentially in a limited time period. A wavelet to be a small wave, it has to satisfy two basic properties:

(i) time integral must be zero

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \quad \text{and} \tag{1}$$

(ii) square of wavelet integrated over time is unity

$$\int_{-\infty}^{\infty} \psi^2(t)dt = 1 \tag{2}$$

Wavelet transform of a 1-D signal $f(x)$ onto a basis of wavelet functions is defined as:

$$W_{a,b}(f(x)) = \int_{x=-\infty}^{\infty} f(x)\psi_{a,b}(x)dx \tag{3}$$

Basis is obtained by translation and dilation of the mother wavelet as:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right) \tag{4}$$

The mother wavelet would localise in both spatial and frequency domain and it has to satisfy zero mean constraint. In discrete wavelet transform (DWT), the dilation factor is $a = 2^m$ and the translation factor is $b = n2^m$, where m and n are integers. The information flow in one level of 2-D image decomposition is illustrated in Fig. 1.

Wavelet separately filters and down samples the 2-D data (image) in the vertical and horizontal directions (separable filter bank). The input (source) image is $I(x,y)$ filtered by low pass filter L and high pass filter H in horizontal direction and then down sampled by a factor of two (keeping the alternative sample) to create the coefficient matrices $I_L(x,y)$ and $I_H(x,y)$ .The coefficient matrices $I_L(x,y)$ and $I_H(x,y)$ are both low pass and high pass filtered in vertical direction and down sampled by

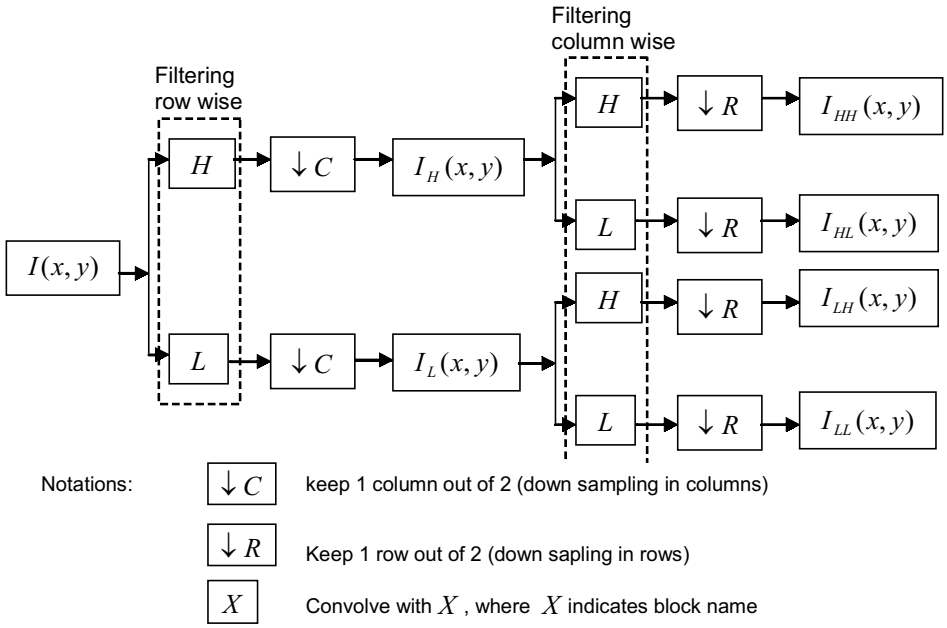


Figure 1. One level of 2-D image decomposition.

a factor of two to create sub bands (sub images) $I_{LL}(x,y)$, $I_{LH}(x,y)$, $I_{HL}(x,y)$, and $I_{HH}(x,y)$ [5].

The $I_{LL}(x,y)$, contains the average image information corresponding to low frequency band of multi scale decomposition. It could be considered as smoothed and sub sampled version of the source image $I(x,y)$. It represents the approximation of source image $I(x,y)$, $I_{LH}(x,y)$, $I_{HL}(x,y)$, and $I_{HH}(x,y)$, are detailed sub images which contain directional (horizontal, vertical and diagonal) information of the source image $I(x,y)$, due to spatial orientation. Multi-resolution could be achieved by recursively applying the same algorithm to low pass coefficients from the previous decomposition[1,5,6].

Inverse 2-D wavelet transform is used to reconstruct the image $I(x,y)$, from sub images $I_{LL}(x,y)$, $I_{LH}(x,y)$, $I_{HL}(x,y)$, and $I_{HH}(x,y)$ as shown in Fig. 2. This involves column up sampling (inserting zeros between samples) and filtering using low pass $\tilde{L}$ and high pass filter $\tilde{H}$ for each sub images. Row up sampling and filtering with low pass filter $\tilde{L}$ and high pass filter $\tilde{H}$ of the resulting image and summation of all matrices would construct the image $I(x,y)$.

2.1.1 Image Fusion by Wavelet Transform

The information flow diagram of wavelet- based image fusion algorithm is shown in Fig. 3. In wavelet image fusion scheme, the source images $I_1(x,y)$ and $I_2(x,y)$, are decomposed into approximation and detailed coefficients at required level using DWT. The approximation and detailed coefficients of both images are combined using fusion rule ϕ. The fused image ($I_f(x,y)$) could be obtained by taking the inverse discrete wavelet transform ($IDWT$) as:

$$I_f(x,y) = IDWT \left[\phi\{DWT(I_1(x,y)), DWT(I_2(x,y))\}\right] \quad (5)$$

The fusion rule used in this paper is simply averages the approximation coefficients and picks the detailed coefficient in each sub band with the largest magnitude.

2.2 Principal Component Analysis

The PCA involves a mathematical procedure that transforms a number of correlated variables into a number of uncorrelated variables called principal components. It computes a compact and optimal description of the data set. The first principal component accounts for as much of the variance in the data as possible and each succeeding component accounts for as much of the remaining variance as possible. First principal component is taken to be along the direction with the maximum variance. The second

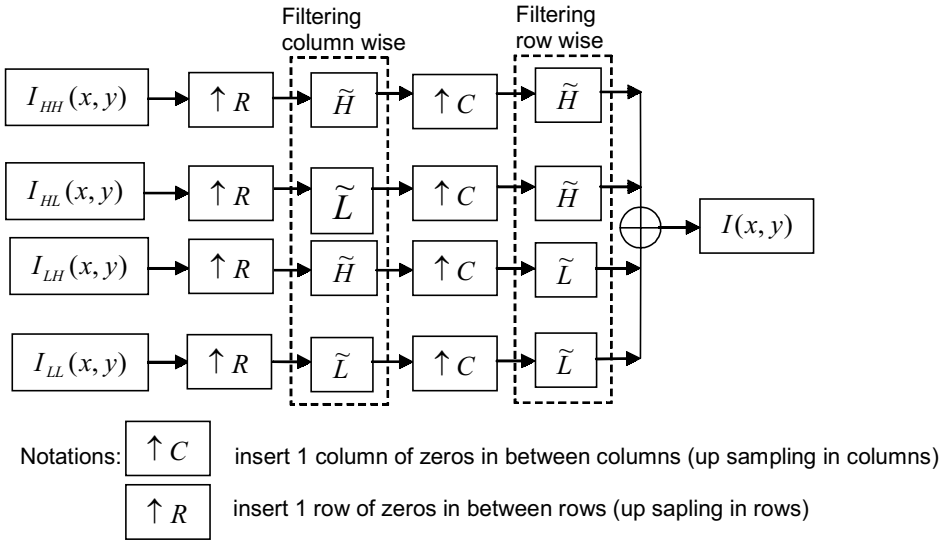


Figure 2. One level of 2-D image reconstruction.

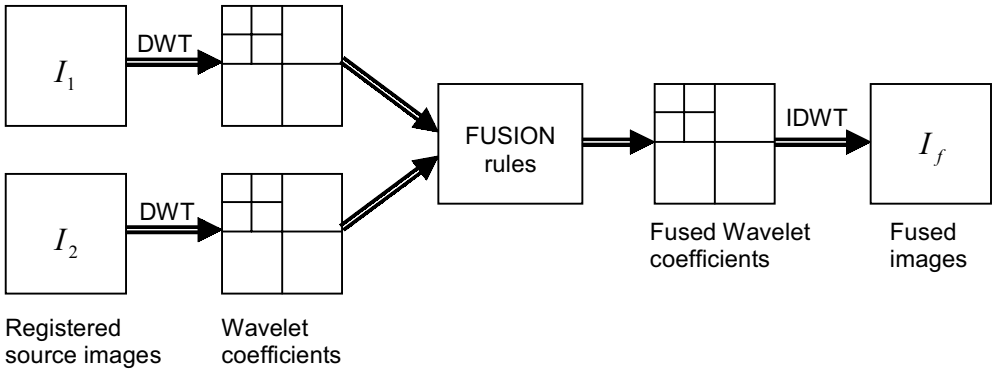


Figure 3. Information flow diagram in image fusion scheme employing multi-scale decomposition.

principal component is constrained to lie in the subspace perpendicular of the first. Within this subspace, this component points the direction of maximum variance. The third principal component is taken in the maximum variance direction in the subspace perpendicular to the first two and so on. The PCA is also called as Karhunen-Loève transform or the Hotelling transform.The PCA does not have a fixed set of basis vectors like FFT, DCT and wavelet etc. and its basis vectors depend on the data set.

Let X be a d-dimensional random vector and assume it to have zero empirical mean. Orthonormal projection matrix V would be such that $Y = V^T X$ with the following constraints. The covariance of Y, i.e., $\text{cov}(Y)$ is a diagonal and inverse of V is equivalent to its transpose ($V^{-1} = V^T$). Using matrix algebra[7]

$$\begin{aligned}\text{cov}(Y) &= E\left\{YY^T\right\} \\ &= E\left\{\left(V^T X\right)\left(V^T X\right)^T\right\} \\ &= E\left\{\left(V^T X\right)\left(X^T V\right)\right\} \\ &= V^T E\left\{XX^T\right\}V \\ &= V^T \text{cov}(X)V\end{aligned} \tag{6}$$

Multiplying both sides of Eqn (6) by V, one gets

$$V\,\text{cov}(Y) = VV^T\,\text{cov}(X)V = \text{cov}(X)V \tag{7}$$

One could write V as $V = [V_1, V_2, ..., V_d]$ and

$$\text{cov}(Y) \text{ as } \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{d-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_d \end{bmatrix} \tag{8}$$

Substituting Eqn (6) into the Eqn (7) gives

$$\begin{aligned}&[\lambda_1 V_1, \lambda_2 V_2, ..., \lambda_d V_d] \\ &= [\text{cov}(X)V_1, \text{cov}(X)V_2, ..., \text{cov}(X)V_d]\end{aligned} \tag{9}$$

This could be rewritten as

$$\lambda_i V_i = \text{cov}(X)V_i \tag{10}$$

where $i = 1, 2, ..., d$ and V_i is an eigenvector of $\text{cov}(X)$.

2.2.1 PCA Algorithm

Let the source images (images to be fused) be arranged in two-column vectors. The steps followed to project this data into a 2-D subspaces are:

1. Organise the data into column vectors. The resulting matrix Z is of dimension 2 x n.
2. Compute the empirical mean along each column. The empirical mean vector M_e has a dimension of 1 x 2.
3. Subtract the empirical mean vector M_e from

each column of the data matrix Z. The resulting matrix X is of dimension 2 x n.

4. Find the covariance matrix C of X i.e.$C=XX_T$ mean of expectation = cov(X)

5. Compute the eigenvectors V and eigenvalue D of C and sort them by decreasing eigenvalue. Both V and D are of dimension 2 x 2.

6. Consider the first column of V which corresponds to larger eigenvalue to compute P_1 and P_2 as:

$$P_1 = \frac{V(1)}{\sum V} \quad \text{and} \quad P_2 = \frac{V(2)}{\sum V} \tag{11}$$

2.2.2 Image Fusion by PCA

The information flow diagram of PCA-based image fusion algorithm is shown in Fig. 4. The input images (images to be fused) $I_1(x,y)$ and $I_2(x,y)$ are arranged in two column vectors and their empirical means are subtracted. The resulting vector has a dimension of n x 2, where n is length of the each image vector. Compute the eigenvector and eigenvalues for this resulting vector are computed and the eigenvectors corresponding to the larger eigenvalue obtained. The normalized components P_1 and P_2 (i.e., $P_1 + P_2 = 1$) using Eqn (9) are computed from the obtained eigenvector. The fused image is:

$$I_f(x,y) = P_1 I_1(x,y) + P_2 I_2(x,y) \tag{12}$$

2.3 Image Fusion by Simple Average

This technique is a basic and straightforward technique and fusion could be achieved by simple averaging corresponding pixels in each input image as:

$$I_f(x,y) = \frac{I_1(x,y) + I_2(x,y)}{2} \tag{13}$$

3. PERFORMANCE EVALUATION

3.1 With Reference Image

When the reference image is available, the performance of image fusion algorithms can be evaluated using the metrics shown in Table 1.

3.2 Without Reference Image

When the reference image is not available, the metrics shown in Table 2 could be used to test the performance of the fused algorithms.

4. RESULTS AND DISCUSSION

Two sets of source/input images are used to evaluate the image fusion algorithms.

4.1 Data Set 1 Analysis

The National Aerospace Laboratories indigenous aircraft *Hansa*, shown in Fig. 5(a) is considered as a reference image to evaluate the performance of the fusion algorithm. The complementary pair

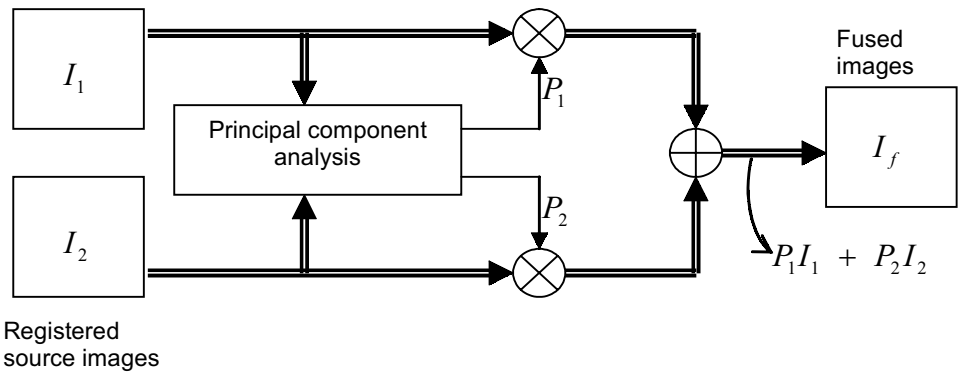


Figure 4. Information flow diagram in image fusion scheme employing PCA.

Table 1. Metrics for performance evaluation when reference image is available

Name of the metric/ ref.	Formula	Features/properties	W fo
Root mean square error[8]	$RMSE = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(I_r(i,j) - I_f(i,j)\right)^2}$	Computed as the root mean square error of the corresponding pixels in the reference image I_r and the fused image I_f.	mi
Mean absolute[8]	$MAE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left\|I_r(i,j) - I_f(i,j)\right\|$	Computed as the mean absolute error of the corresponding pixels in reference and fused images.	mi
Percentage fit error[8]	$PFE = \frac{norm(I_r - I_f)}{norm(I_r)} * 100$ where $norm$ is the operator to compute the largest singular value	Computed as the norm of the difference between the corresponding pixels of reference and fused image to the norm of the reference image. This will be zero when both reference and fused images are exactly alike and it will be increased when the fused image is deviated from the reference image.	M
Signal to noise ratio[9]	$SNR = 20\log_{10}\left(\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(I_r(i,j)\right)^2}{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(I_r(i,j) - I_f(i,j)\right)^2}\right)$	This metric will be high when the reference and fused images are alike. Higher value implies better fusion.	ma
Peak signal to noise[9]	$PSNR = 20\log_{10}\left(\frac{L^2}{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(I_r(i,j) - I_f(i,j)\right)^2}\right)$ where L in the number of gray levels in the image	Its value will be high when the fused and reference images are similar. Higher value implies better fusion.	ma
Correlation[8]	$CORR = \frac{2C_{rf}}{C_r + C_f}$ where $C_r = \sum_{i=1}^{M}\sum_{j=1}^{N} I_r(i,j)^2$ $C_f = \sum_{i=1}^{M}\sum_{j=1}^{N} I_f(i,j)^2$ and $C_{rf} = \sum_{i=1}^{M}\sum_{j=1}^{N} I_r(i,j) I_f(i,j)$	This shows the correlation between the reference and fused image. The ideal value is one when the reference and fused are exactly alike and it will be less than one when the dissimilarity increases.	ma
Mutual information[11]	$MI = \sum_{i=1}^{M}\sum_{j=1}^{N} h_{I_r I_f}(i,j)\log_2\left(\frac{h_{I_r I_f}(i,j)}{h_{I_r}(i,j) h_{I_f}(i,j)}\right)$	Larger value implies better image quality.	ma
Universal quality index[12]	$QI = \frac{4\sigma_{I_r I_f}(\mu_{I_r} + \mu_{I_f})}{(\sigma_{I_r}^2 + \sigma_{I_f}^2)(\mu_{I_r}^2 + \mu_{I_f}^2)}$ where $\mu_{I_r} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} I_r(i,j)$, $\mu_{I_f} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} I_f(i,j)$, $\sigma_{I_r}^2 = \frac{1}{MN-1}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(I_r(i,j) - \mu_{I_r}\right)^2$, $\sigma_{I_f}^2 = \frac{1}{MN-1}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(I_f(i,j) - \mu_{I_f}\right)^2$, $\sigma_{I_r I_f}^2 = \frac{1}{MN-1}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(I_r(i,j) - \mu_{I_r}\right)\left(I_f(i,j) - \mu_{I_f}\right)$	This measures how much of the salient information contained in reference image has been transformed into the fused image. The range of this metric is -1 to 1 and the best value 1 would be achieved if and only if reference and fused images are alike. The lowest value of -1 would occur when $I_f = 2\mu_{I_r} - I_r$.	ma
Measure of structural similarity[10]	$SSIM = \frac{(2\mu_{I_r}\mu_{I_f} + C_1)(2\sigma_{I_r I_f} + C_2)}{(\mu_{I_r}^2 + \mu_{I_f}^2 + C_1)(\sigma_{I_r}^2 + \sigma_{I_f}^2 + C_2)}$ where C_1 is a constant that is included to avoid the instability when $\mu_{I_r}^2 + \mu_{I_f}^2$ is close to zero and C_2 is a constant that is included to avoid the instability when $\sigma_{I_r}^2 + \sigma_{I_f}^2$ is close to zero	Natural image signals would be highly structured and their pixels reveal strong dependencies. These dependencies would carry vital information about the structure of the object. It compares local patterns of pixel intensities that have been normalized for luminance and contrast.	ma

Table 2. Metrics for performance evaluation when reference image is not

Name of the metric	Formula	Features/properties	What look f min/m
Standard deviation[10]	$\sigma = \sqrt{\sum_{i=0}^{L}(i-\bar{i})^2 h_{I_f}(i)}$, $\bar{i} = \sum_{i=0}^{L} i h_{I_f}$ where $h_{I_f}(i)$ is the normalized histogram of the fused image $I_f(x,y)$ and L number of frequency bins in the histogram.	It is known that standard deviation is composed of the signal and noise parts. This metric would be more efficient in the absence of noise. It measures the contrast in the fused image. An image with high contrast would have a high standard deviation.	maxin
Entropy[13]	Using the entropy, the information content of a fused image is: $He = -\sum_{i=0}^{L} h_{I_f}(i)\log_2 h_{I_f}(i)$	Entropy is used to measure the information content of an image. Entropy is sensitive to noise and other unwanted rapid fluctuations. An image with high information content would have high entropy.	maxin
Cross entropy[13]	overall cross entropy of the source images I_1, I_2 and the fused image I_f is: $CE(I_1,I_2;I_f) = \frac{CE(I_1;I_f)+CE(I_2;I_f)}{2}$ where $CE(I_1;I_f) = \sum_{i=0}^{L} h_{I_1}(i)\log\left(\frac{h_{I_1}(i)}{h_{I_f}(i)}\right)$ & $CE(I_2;I_f) = \sum_{i=0}^{L} h_{I_2}(i)\log\left(\frac{h_{I_2}(i)}{h_{I_f}(i)}\right)$	Cross-entropy evaluates the similarity in information content between input images and fused image. Fused and reference images containing the same information would have a low cross entropy.	minim
Spatial frequency[14]	spatial frequency criterion SF is: $SF = \sqrt{RF^2 + CF^2}$ where row frequency of the image: $RF = \sqrt{\frac{1}{MN}\sum_{x=1}^{M}\sum_{y=2}^{N}[I_f(x,y)-I_f(x,y-1)]^2}$ column frequency of the image: $CF = \sqrt{\frac{1}{MN}\sum_{y=1}^{N}\sum_{x=2}^{M}[I_f(x,y)-I_f(x-1,y)]^2}$	This frequency in spatial domain indicates the overall activity level in the fused image.	maxin
Fusion mutual information[15]	If the joint histogram between $I_1(x,y)$ and $I_f(x,y)$ is defined as $h_{I_1I_f}(i,j)$ and $I_2(x,y)$ and $I_f(x,y)$ as $h_{I_2I_f}(i,j)$. Then the mutual information between source and fused images are: $FMI = MI_{I_1I_f} + MI_{I_2I_f}$ where $MI_{I_1I_f} = \sum_{i=1}^{M}\sum_{j=1}^{N} h_{I_1I_f}(i,j)\log_2\left(\frac{h_{I_1I_f}(i,j)}{h_{I_1}(i,j)h_{I_f}(i,j)}\right)$ $MI_{I_2I_f} = \sum_{i=1}^{M}\sum_{j=1}^{N} h_{I_2I_f}(i,j)\log_2\left(\frac{h_{I_2I_f}(i,j)}{h_{I_2}(i,j)h_{I_f}(i,j)}\right)$	It measures the degree of dependence of the two images. A larger measure implies better quality.	maxin
Fusion quality index[16]	$FQI = \sum_{w\in W} c(w)\left(\lambda(w)QI(I_1,I_f \mid w) + (1-\lambda(w))QI(I_2,I_f \mid w)\right)$ where $\lambda(w) = \frac{\sigma_{I_1}^2}{\sigma_{I_1}^2+\sigma_{I_2}^2}$ computed over a window; $C(w) = \max\left(\sigma_{I_1}^2, \sigma_{I_2}^2\right)$ over a window $c(w)$ is a normalized version of $C(w)$ & $QI(I_1,I_f \mid w)$ is the quality index over a window for a given source image and fused image	The range of this metric is 0 to 1. One indicates the fused image contains all the information from the source images.	maxin
Fusion similarity metric[15]	$FSM = \sum_{w\in W} sim(I_1,I_2,I_f \mid w)\begin{pmatrix} QI(I_1,I_f \mid w) \\ -QI(I_2,I_f \mid w)\end{pmatrix} + QI(I_2,I_f \mid w)$ where $sim(I_1,I_2,I_f \mid w) = \begin{cases} 0 & \text{if } \frac{\sigma_{I_1I_f}}{\sigma_{I_1I_f}+\sigma_{I_2I_f}} < 0 \\ \frac{\sigma_{I_1I_f}}{\sigma_{I_1I_f}+\sigma_{I_2I_f}} & \text{if } 0 \le \frac{\sigma_{I_1I_f}}{\sigma_{I_1I_f}+\sigma_{I_2I_f}} \le 1 \\ 1 & \text{if } \frac{\sigma_{I_1I_f}}{\sigma_{I_1I_f}+\sigma_{I_2I_f}} > 1 \end{cases}$	It takes into account the similarity between the source and fused image block within the same spatial position. The range of this metric is zero to one. The value one indicates that the fused image contains all the information from the source images.	maxin

input images and are taken to evaluate the fusion algorithm and these images are shown in Fig. 5(b) and 5(c). The complementary pair has been created by blurring the reference image of size 295 x 400 with a Gaussian mask using diameter of 12 pixels. The images are complementary in the sense that the blurring occurs at the left-half and the right-half respectively. The first column in Figs 6 to 10 show fused images and the second column shows the error images. The error (difference) image is computed by taking the corresponding pixel difference of reference mage and fused image, i.e., $I_e(x,y) = I_r(x,y) - I_f(x,y)$.

(a) Reference image $I_r(x, y)$

(b) Source image $I_1(x, y)$

(c) Source image $I_2(x, y)$

Figure 5. Reference and source images (data set 1).

The fused and error images by simple average image fusion algorithm are shown in Fig. 6. Similarly, the fused and error images by PCA algorithm are shown in Fig. 7. It is observed that the performance of average and PCA fusion algorithms are similar for these images. The reason could be because of taking the complementary pairs. The principal components are [0.5069, 0.4931], is approximately 0.5, and in this situation, the PCA is equivalent to a simply average [Eqn (11)]. The fused and error images by wavelet with different levels of decomposition are shown in Figs 8 to 10. First two letters indicate the wavelet and the number following the letters indicates the level of decomposition.

(a) Fused image

(b) Error image

Figure 6. Fused and error images by simple average (data set 1).

By observing the error images from Figs 8 to 10, it is seen that better fusion could be achieved with

(a) Fused image

(b) Error image

Figure 7. Fused and error images by PCA algorithm (data set 1).

(a) Fused image

(b) Error image

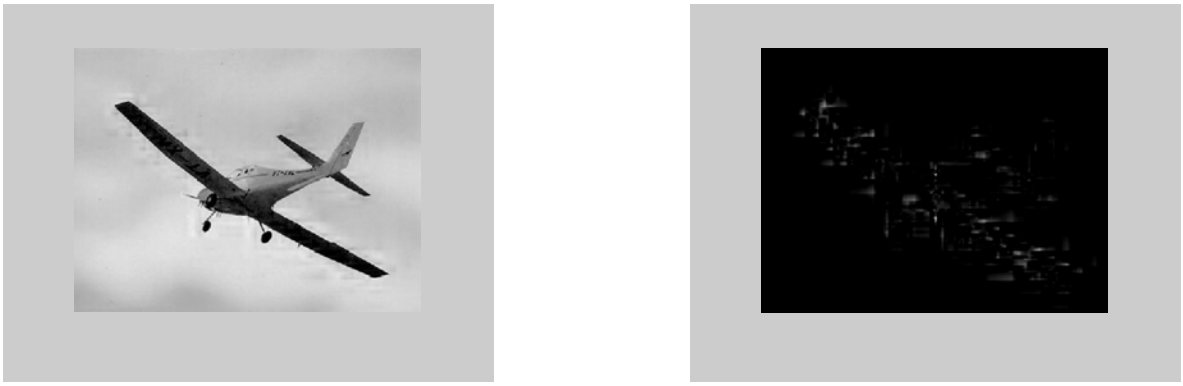
Figure 8. Fused and error images by wavelet 1 (wt1) (data set 1).

high level of decomposition. It is possibly due to consideration of all frequency bands in the process of fusion. From Figs 8 to 10, it is also observed that there are increased oscillations (ringing tone) presented in the fused image at the sharp edges. These oscillations are high when wavelets with higher level of decomposition are used in fusion process. It is similar to Gibb's phenomenon. This ringing tone could be avoided using wavelets with shift invariant property. The performance metrics for evaluating the image fusion algorithms are shown in Table 3. Other performance metrics where the reference image is not considered are shown in Table 4. The metrics showed in tables in bold font are better among others. From the tables it is observed that except in few metrics, wavelet transform with higher level of decomposition performed well.

(a) Fused image

(b) Error image

Figure 9. Fused and error images by wavelet 3 (wt3) (data set 1).



(a) Fused image (b) Error image

Figure 10. Fused and error images by wavelet 5 (wt5) (data set 1).

For this data set, separate fusion algorithms were actually run for each of the left and right half of the picture and it was found that the results were very close to those of Tables 3 and 4. Hence the new results are not included in this paper.

4.2 Data Set 2 Analysis

In this data set, forward-looking infrared image and low-light television image of size 512 x 512 are considered for evaluation of the fusion algorithms. These images are shown in Fig. 11. Reference image is not available for this data set. The fused image obtained by simple average and PCA are shown in Figs 12 and 13 respectively. Figures 14 to 18 show the fused images by wavelets with different level of decomposition. The performance metrics, where the reference image is not available, are shown in Table 5. It is observed that simple average image fusion algorithm shows degraded performance. It could be due to combination of source images taken from different modalities/sensors. It is observed that in some metrics, PCA shows better performance and the other metrics, wavelets with higher level of decomposition show better results.

5. OBSERVATIONS

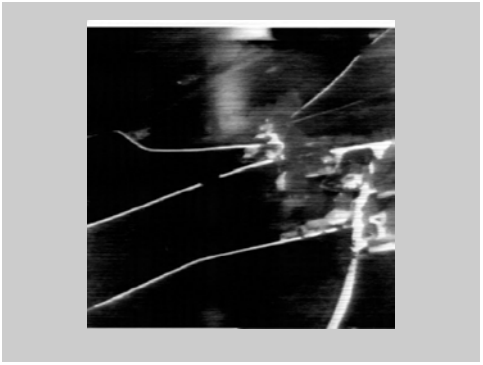
The performance of the image fusion algorithms could be estimated precisely when the ground truth (reference) image is available. The reference image is available in data set 1 and the performance metrics are shown in Table 3. It is observed that wavelets with higher level of decomposition show

Table 3. Performance evaluation metrics to evaluate image fusion algorithms with reference images (data set 1)

	RMSE	PFE	MAE	CORR	SNR	PSNR	MI	QI	SSIM
Ave	9.91	5.11	5.18	0.9987	51.66	76.42	1.40	0.81	0.95
PCA	9.86	5.09	5.15	0.9987	51.74	76.44	1.40	0.81	0.95
wt1	9.18	4.73	4.90	0.9985	53.00	77.08	1.41	0.85	0.96
wt2	8.05	4.15	4.50	0.9991	55.26	78.20	1.43	0.88	0.96
wt3	6.49	3.35	3.92	0.9994	59.00	80.08	1.44	0.88	0.96
wt4	4.79	2.47	3.32	0.9997	64.30	82.72	1.41	0.86	0.97
wt5	3.96	2.04	3.00	0.9998	67.58	84.36	1.40	0.85	0.97

Table 4. Performance evaluation metrics to evaluate image fusion algorithms without reference image (data set 1)

	He	SD	CE	SF	FMI	FQI	FSM
Ave	3.4	45.23	0.53	10.91	2.79	0.83	0.78
PCA	3.41	45.23	0.51	10.94	2.79	0.83	0.78
wt1	3.31	45.72	0.15	16.79	2.75	0.82	0.75
wt2	3.29	46.43	0.31	19.11	2.73	0.81	0.73
wt3	3.24	47.50	0.83	19.87	2.71	0.79	0.7
wt4	3.25	48.95	0.27	20.10	2.65	0.77	0.66
wt5	3.26	50.25	0.11	20.16	2.62	0.77	0.66



(a) FLIR image $I_1(x, y)$

(b) TV image $I_2(x, y)$

Figure 11. Source images used for image fusion (data set 2).

superior results except in mutual information (MI) and universal quality index (QI). Some metrics standard deviatio (SD), CE and spatial frequency (SF) in Table 4 where the reference image is not considered in estimation also showed the similar observation. Other metrics (H),fusion mutual information (FMI), fusion quality index (FQI) and fusion similarity matrix (FSM) failed to give similar observation. From these results one can observe that the metrics SD, CE and SF would be appropriate metrics to access the quality of the fused image when there is no reference image is available.

The metrics SD, CE, and SF show different results for data set 2 where the images to be fused are obtained from different modalities as shown in Table 5. The SD showed that PCA is better but by observing the fused image, it is not. It may be

Figure 12. Fused image by simple average.

Figure 13. Fused image by PCA (data set 2).

Figure 14. Fused image by wavelet transform (wt1) (data set 2).

Figure 15. Fused image by wavelet transform (wt2).

Table 5. Performance evaluation metrics to evaluate image fusion algorithms without reference image (data set 2)

	He	SD	CE	SF	FMI	FQI	FSM
Ave	7.46	47.39	2.37	10.15	2.27	0.56	0.38
PCA	7.87	96.47	2.64	13.52	2.44	0.54	0.40
wt1	7.30	47.83	4.38	14.97	2.19	0.59	0.42
wt2	7.19	48.56	5.03	17.11	2.18	0.60	0.44
wt3	6.96	50.29	5.71	18.54	2.16	0.61	0.45
wt4	6.84	54.49	5.84	19.37	2.15	0.61	0.45
wt5	6.91	58.30	5.77	19.62	2.14	0.60	0.43

Figure 16. Fused image by wavelet transform (wt3) (data set 2).

Figure 17. Fused image by wavelet transform (wt4).

Figure 18. Fused image by wavelet transform (wt5) (data set 2).

due to the consideration of contrast in SD calculations. From Table 4 and Table 5, it is observed that SD would be a better fusion indicator where the images to be fused are obtained from the same or similar modalities. The other metric SF showed that wavelets with higher decomposition level show better performance. From Table 4 and 5, it is observed that SF would be the best fusion performance check indicator irrespective of the origin of source images.

6. CONCLUSIONS

Pixel-level image fusion using wavelet transform and principal component analysis are implemented in PC MATLAB. Different image fusion performance metrics with and without reference image have been evaluated. The simple averaging fusion algorithm shows degraded performance. Image fusion using wavelets with higher level of decomposition shows better performance in some metrics while in other metrics, the PCA shows better performance. Some further investigation is needed to resolve this issue.

ACKNOWLEDGEMENTS

The authors would like to thank Dr Girija G. and Dr Jatinder Singh, Scientists, Flight Mechanics and Control Division, National Aerospace Laboratories, Bangalore, for technical discussions.

REFERENCES

1. Gonzalo, Pajares & Jesus Manuel, de la Cruz. A wavelet-based image fusion tutorial. *Pattern Recognition*, 2004, **37**, 1855-872.

2. Varsheny, P.K. Multi-sensor data fusion. *Elec. Comm. Engg.*, 1997, **9**(12), 245-53.

3. Mallet, S.G. A heory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intel.*, 1989, **11**(7), 674-93.

4. Wang, H.; Peng, J. & Wu, W. Fusion algorithm for multisensor image based on discrete multiwavelet transform. *IEE Proc. Visual Image Signal Process.*, 2002, **149**(5).

5. Mitra Jalili-Moghaddam. Real-time multi-focus image fusion using discrete wavelet transform and Laplasican pyramid transform. Chalmess University of Technology, Goteborg, Sweden, 2005. Masters thesis.

6. Daubechies, I. Ten lectures on wavelets. *In* Regular Conference Series in Applied Maths, Vol. 91, 1992, SIAM, Philadelphia.

7. http://en.wikipedia.org/wiki/Principal_components_analysis

8. Naidu, V.P.S.; Girija, G. & Raol, J.R. Evaluation of data association and fusion algorithms for tracking in the presence of measurement loss.*In* AIAA Conference on Navigation, Guidance and Control, Austin, USA, August 2003, pp. 11-14.

9. Arce, Gonzalo R. Nonlinear signal processing A statistical approach. Wiley-Interscience Inc. Publication, USA, 2005.

10. Blum, Rick S. & Zheng, Liu. Multi-sensor image fusion and its applications. CRC Press, Taylor & Francis Group, Boca Raton, 2006.

11. Cover, T.M. & Thomas, J.A. Elements of information theory. Wiley, New York, 1991.

12. Wang, Z. & Bovik, A.C. A universal image quality index. *IEEE Signal Proc. Letters*, March 2002, **9**(9), 81-84.

13. Leung, lau Wai.; King, Bruce & Vohora, Vijay. Comparison of image data fusion techniques using entropy and INI. *In* 22nd Asian Conference on Remote Sensing, Singapore, Nov. 2001, pp. 5-9.

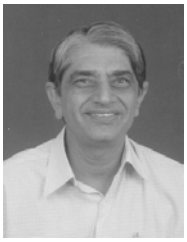
14. Eskicioglu, A.M. & Fisher, P.S. Image quality measures and their performance. *IEEE Trans. Commu.,* 1995, **43**(12), 2959-965.

15. Cvejic, Nedeljko; Loza, Artur. ; Bull, David & Cangarajah, Nishan. A similarity metric for assessment of image fusion algorithms. *Int. J. Signal Process.*, 2005, **2**(3), 178-82.

16. Piella, Gemma & Heijmans, Henk. A new quality metric for image fusion. *In* Proceedings of IEEE International Conference on Image Processing, Barcelona, Spain, 2003, pp. 173-76.

Contributors

Mr V.P.S. Naidu obtained his ME (Medical Electronics) from Anna University, Chennai, in 1997. He is working as a Scientist at National Aerospace Laboratories (NAL), Bangalore, since December 2001 in the area of multi-sensor data fusion and target tracking. His areas of interest are image registration, tracking, and data fusion.



Dr J.R. Raol obtained BE and ME, both from the MS University, Baroda, Vadodara, in 1971 and 1974 and PhD from McMaster University, Canada, in 1986. He worked at NAL Bangalore from 1975 to 1981 and was actively involved in the multidisciplinary control group's activities on human pilot modelling in fix- and motion-based research simulators. He re-joined NAL in 1986 and served as a senior scientist and the Head of the Flight Mechanics and Control Division for several years. He has authored jointly a book on "Modelling and Parameter Estimation of dynamic system", published by IEE, UK, in 2004. He has 100 research publications to his credit. He has carried out sponsored R&D in several areas. His current activities include: Modelling, parameter estimation, multi-sensor data fusion, fuzzy systems, genetic algorithms and neural networks. Currently he is the Professor Emeritus in the Dept of Instrumentation Technology of the MS Ramaiah Institute of Technology, Bangalore.